# Dependence of cosmic shear covariances on cosmology

## Impact on parameter estimation

T. Eifler[1], P. Schneider[1] and J. Hartlap[1]

1) Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany

**ABSTRACT**

*Context.* In cosmic shear likelihood analyses the covariance is most commonly assumed to be constant in parameter space. Therefore, when calculating the covariance matrix (analytically or from simulations), its underlying cosmology should not influence the likelihood contours.
*Aims.* We examine whether the aforementioned assumptions hold and quantify how strong cosmic shear covariances vary within a reasonable parameter range. Furthermore, we examine the impact on likelihood contours when assuming different cosmologies in the covariance. The final goal is to develop an improved likelihood analysis for parameter estimation with cosmic shear.
*Methods.* We calculate Gaussian covariances analytically for 2500 different cosmologies. In order to quantify the impact on the parameter constraints we perform a likelihood analysis for each covariance matrix and compare the likelihood contours. To improve on the assumption of a constant covariance, we use an adaptive covariance matrix, which is continuously updated according to the point in parameter space where the likelihood is evaluated. As a side-effect, this cosmology dependent covariance improves the parameter constraints. We examine this fact more closely using the Fisher-matrix formalism. In addition we quantify the impact of non-Gaussian covariances on the likelihood contours using a ray-tracing covariance derived from the Millennium simulation. In this ansatz we return to the approximation of a constant covariance matrix; in order to minimize the error due to this approximation, we develop the concept of an iterative likelihood analysis.
*Results.* Covariances vary significantly within the considered parameter range. The cosmology assumed in the covariance has a non-negligible impact on the size of the likelihood contours. This impact increases with increasing survey size, increasing number density of source galaxies, decreasing ellipticity noise, and when taking non-Gaussianity into account. A proper treatment of this effect is therefore even more important for future surveys. In this paper we present methods to take cosmology dependent covariances into account.

**Key words.** cosmology: theory - gravitational lensing - large-scale structure of the Universe - methods: statistical

## 1. Introduction

Cosmic shear, which was first detected in 2000 (Bacon et al. 2000; Kaiser et al. 2000; van Waerbeke et al. 2000; Wittman et al. 2000), has recently progressed to an important tool in cosmology. Latest results (e.g. van Waerbeke et al. 2005; Semboloni et al. 2006; Hoekstra et al. 2006; Schrabback et al. 2007; Hetterscheidt et al. 2007; Massey et al. 2007; Fu et al. 2008) already indicate its great ability to constrain cosmological parameters which will be enhanced in the future by large upcoming surveys like Pan-STARRS, KIDS, DES, Euclid or LSST. The improved quality of cosmic shear data must be accompanied with an accurate data analysis, free of assumptions which bias the results. Obtaining appropriate covariances is a crucial issue in this context of a precision cosmology likelihood analysis. Several methods are suggested in the literature and have been applied to cosmic shear data. An analytic expression for covariances assuming a Gaussian shear field is derived in Schneider et al. (2002a) and confirmed

in Joachimi et al. (2008) who use a power spectrum approach which significantly reduces the computational effort in the calculation. This analytic expression has been used for parameter estimation in many surveys (e.g. van Waerbeke et al. 2005; Semboloni et al. 2006; Hoekstra et al. 2006). However, the assumption of a Gaussian shear field breaks down on small scales; according to Kilbinger & Schneider (2005) and Semboloni et al. (2007) non-linear effects already become important at angular scales $\lesssim 10$ arcmin. To account for non-Gaussianity, Semboloni et al. (2007) invent a calibration factor which is derived from a comparison of Gaussian to ray-tracing covariances. An application of this method to real data can be found in (Fu et al. 2008). A second approach is the derivation of the covariance matrix from the data (e.g. Hetterscheidt et al. 2007; Massey et al. 2007). Here, the covariance is calculated via field-to-field variation which involves a separation of the data set into many independent subsamples. This might lead to a loss of information on large scales if the survey is not sufficiently large. Third, one can estimate the covariance matrix from ray-tracing simulations, a method which circumvents the

aforementioned loss in information. Although, in this method the covariance is again derived via field-to-field variation, we can choose a sufficiently large numerical simulation to create many independent subsamples of adequate size.

Note, that the last two methods involve an estimation process in the determination of the covariance matrix, which means that the inverse is biased and one has to correct for this effect (Anderson 2003; Hartlap et al. 2007). Nevertheless, deriving covariance matrices from ray-tracing simulations seems to be a promising method as it preserves all the information in the data and additionally takes the non-Gaussianity of the shear field into account.

The analytic expression and the ray-tracing covariance assume a specific cosmological model in their derivation. So far, cosmic shear likelihood analyses treat the covariance matrix as constant in parameter space, hence its underlying cosmology is assumed not to influence the parameter constraints. It is the intention of this paper to check for this assumption and in case it does not hold, to present an improved likelihood formalism for future surveys.

This paper is organized as follows. Section 2 summarizes the basic theoretical background of the cosmic shear two-point correlation function (2PCF) and its corresponding covariance. In Sect. 3 we derive a scaling relation for covariances, which can be used for a fast calculation of covariances for arbitrary cosmology. Furthermore, we examine how strongly the covariance depends on its underlying cosmological model. The impact on parameter constraints when assuming a fixed cosmology in the covariance is subject of Sect. 4, whereas we present improvements on this assumption in Sect. 5. Here, we consider a likelihood analysis with an adaptive covariance matrix, i.e. the covariance is calculated individually for each point in parameter space where the likelihood is evaluated. In addition, we outline the concept of an iterative covariance matrix, i.e. several likelihood analyses are performed, where the covariance is updated in every iteration according to the maximum likelihood parameter set. Here, we also examine the impact of non-Gaussian covariances on the likelihood contours using a ray-tracing covariance matrix derived from the Millennium Simulation. We conclude in Sect. 6.

## 2. Data vectors and covariances of cosmic shear

In this section we briefly review the basics of the cosmic shear two-point correlation function and its corresponding covariance matrix. For more details on this topic the reader is referred to Bartelmann & Schneider (2001); Schneider et al. (2002a,b); Kilbinger & Schneider (2004); Joachimi et al. (2008).

To measure the shear signal we define $\boldsymbol{\theta}$ as the connecting vector of two points and specify tangential and cross-component of the shear $\gamma$ as

$$\gamma_t = -\mathrm{Re}\left(\gamma e^{-2i\varphi}\right) \qquad \text{and} \qquad \gamma_\times = -\mathrm{Im}\left(\gamma e^{-2i\varphi}\right), \tag{1}$$

where $\varphi$ is the polar angle of $\boldsymbol{\theta}$. The 2PCFs depend only on the absolute value of $\boldsymbol{\theta}$. They are defined in terms of the shear and can be related to the power spectra $P_\mathrm{E}$ and $P_\mathrm{B}$ (Schneider et al. 2002b)

$$\xi_\pm(\theta) \equiv \langle \gamma_t \gamma_t \rangle(\theta) \pm \langle \gamma_\times \gamma_\times \rangle(\theta) \tag{2}$$

$$= \int_0^\infty \frac{d\ell\,\ell}{2\pi}\,\mathrm{J}_{0/4}(\ell\theta)\,\left[P_\mathrm{E}(\ell) \pm P_\mathrm{B}(\ell)\right], \tag{3}$$

with $\mathrm{J}_n$ denoting the $n$-th order Bessel-function. In this paper we only consider E-modes, therefore we set $P_\mathrm{B} = 0$ from now on. Furthermore, we assume that the 2PCF is estimated in logarithmic bins $\vartheta$ of angular width $\Delta\vartheta$. The covariance of the 2PCF is defined as

$$C_\xi\left(\vartheta_i, \vartheta_j\right) := \left\langle \left(\xi_\pm(\vartheta_i) - \hat{\xi}_\pm(\vartheta_i)\right)\left(\xi_\pm(\vartheta_j) - \hat{\xi}_\pm(\vartheta_j)\right)\right\rangle. \tag{4}$$

We neglect the index $\xi$ in the covariance for the rest of the paper as we only consider covariances of the 2PCF. As one already sees from (4) the 2PCF has four different covariances, denoted as $C_{++}$, $C_{+-}$, $C_{-+}$, $C_{--}$. Only three of them are independent since $C_{+-}(\vartheta_i, \vartheta_j) = C_{-+}(\vartheta_j, \vartheta_i)$. Assuming a Gaussian shear field the covariance of the 2PCF can be calculated analytically (Schneider et al. 2002a; Joachimi et al. 2008). There, the covariance is decomposed into three terms, namely the cosmic variance term (V), the pure shot noise term (S), and the mixed term (M)

$$C_{++}(\vartheta_i, \vartheta_j) = V_{++} + M_{++} + S, \tag{5}$$
$$C_{--}(\vartheta_i, \vartheta_j) = V_{--} + M_{--} + S, \tag{6}$$
$$C_{+-}(\vartheta_i, \vartheta_j) = V_{+-} + M_{+-}. \tag{7}$$

The pure shot noise term vanishes in case of $C_{+-}$ and only contributes to the diagonal of $C_{++}$ and $C_{--}$. It can be calculated as

$$S = \frac{\sigma_\epsilon^4}{2\pi\vartheta_i\Delta\vartheta_i A\bar{n}^2}\,\delta_{\vartheta_i\vartheta_j}, \tag{8}$$

where $A$ denotes the solid angle of the data field, $\sigma_\epsilon$ is the intrinsic ellipticity dispersion, and $\bar{n}$ the number density of source galaxies. The cosmic variance term (V) and the mixed term (M) can be either calculated using the power spectrum or the 2PCF. According to Joachimi et al. (2008) the power spectrum approach leads to the following expressions

$$V_{\pm\pm} = \frac{1}{\pi A} \int_0^\infty d\ell\,\ell\,\mathrm{J}_{0/4}(\ell\vartheta_i)\,\mathrm{J}_{0/4}(\ell\vartheta_j)\,P_\mathrm{E}^2(\ell), \tag{9}$$

$$M_{\pm\pm} = \frac{\sigma_\epsilon^2}{\pi A\bar{n}} \int_0^\infty d\ell\,\ell\,\mathrm{J}_{0/4}(\ell\vartheta_i)\,\mathrm{J}_{0/4}(\ell\vartheta_j)\,P_\mathrm{E}(\ell). \tag{10}$$

The corresponding expressions for $V$ and $M$ using the 2PCF are derived in Schneider et al. (2002a). In this paper we only need the expressions for the mixed term, which read

$$M_{++} = \frac{2\sigma_\epsilon^2}{\pi A n} \int_0^\pi d\varphi\,\xi_+(|\boldsymbol{\phi}|), \tag{11}$$

$$M_{--} = \frac{2\sigma_\epsilon^2}{\pi A n} \int_0^\pi d\varphi\,\xi_+(|\boldsymbol{\phi}|)\,\cos(4\varphi), \tag{12}$$

$$M_{+-} = \frac{2\sigma_\epsilon^2}{\pi A n} \int_0^\pi d\varphi\left[\sum_{k=0}^4 \binom{4}{k}(-1)^k \vartheta_i^k\,\vartheta_j^{4-k}\cos(k\varphi)\right]$$
$$\times (|\boldsymbol{\phi}|)^{-4}\,\xi_-(|\boldsymbol{\phi}|)\,\cos(4\varphi), \tag{13}$$

where we denote $|\boldsymbol{\phi}| = \sqrt{\vartheta_i^2 + \vartheta_j^2 - 2\vartheta_i\vartheta_j\cos\varphi}$.

## 3. Variation of covariances in parameter space

We select a two-dimensional parameter grid with $50 \times 50$ gridpoints of $\Omega_{\mathrm{m}} \in [0.2; 0.4]$ and $\sigma_8 \in [0.6; 1.0]$. For each grid point we calculate a covariance analytically using (5) - (10). The shear power spectra $P_{\mathrm{E}}$ are obtained from the density power spectra $P_{\delta}$ employing Limber's equation. To derive $P_{\delta}$ we assume an initial Harrison-Zeldovich power spectrum ($P_{\delta}(k) \propto k^{n_{\mathrm{s}}}$ where $n_{\mathrm{s}} = 1$) with the transfer function from Efstathiou et al. (1992). For the calculation of the non-linear evolution we use the fitting formula of Smith et al. (2003). Throughout this paper, we assume a flat universe and fix all cosmological parameters except $\Omega_{\mathrm{m}}$ and $\sigma_8$, more precisely $H_0 = 0.73$ and $\Omega_{\mathrm{b}} = 0.04$. These values for $H_0$ and $\Omega_{\mathrm{b}}$ together with $\Omega_{\mathrm{m}} = 0.25$ and $\sigma_8 = 0.9$ define our fiducial cosmological model, which we have chosen similar to the cosmology of the Millennium Simulation (Springel et al. 2005) for a later comparison of Gaussian and ray-tracing covariances. We assume all source galaxies to be at redshift $z_0 = 1.0$. Using a redshift distribution instead would not change our results qualitatively. In addition to cosmology, the covariance depends on survey parameters. The scaling relations given in Sect. 3 are generally valid and independent of survey parameters. In case of the likelihood analyses in Sects. 4 and 5 we choose, unless stated otherwise, an intrinsic ellipticity noise of $\sigma_{\epsilon} = 0.4$, a number density of source galaxies of $\bar{n} = 10/\mathrm{arcmin}^2$ (similar to the values of the Dark Energy survey), and a survey which covers $A = 900$ deg$^2$. The angular scale of the 2PCF data vector for which we calculate the covariances covers a range from 0.1 arcmin to 180 arcmin, which is divided into 50 logarithmic bins.

### 3.1. A fast method to calculate covariances for arbitrary $\Omega_{\mathrm{m}}$ and $\sigma_8$

From (9) and (10) one directly sees that the covariance matrix depends on the cosmological model, which enters with the power spectrum $P_{\mathrm{E}}$. Figure 1 illustrates the change in $P_{\mathrm{E}}$ when varying only $\Omega_{\mathrm{m}}$, or $\sigma_8$, and both parameters simultaneously; we see that it increases with $\Omega_{\mathrm{m}}$ as well as with $\sigma_8$.

For a given cosmological model we can calculate the covariance directly from (5) - (10). Performing this calculation for many sets of parameters is time-consuming; hence we seek a scaling relation, which relates the covariances of an arbitrary cosmology $\mathbf{C}_{\pi}$ to a reference model $\mathbf{C}_{\pi_0}$. A basic theorem in statistics states (e.g. Anderson 2003), that if there is a relation between two data vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ which reads $\boldsymbol{y} = \mathbf{A}\boldsymbol{x}$ ($\mathbf{A}$ being a matrix), the relation of the covariances of $\boldsymbol{x}$ and $\boldsymbol{y}$ can be written as

$$\mathbf{C}_y = \left\langle (\boldsymbol{y} - \langle \boldsymbol{y} \rangle)(\boldsymbol{y} - \langle \boldsymbol{y} \rangle)^{\mathrm{t}} \right\rangle$$
$$= \left\langle (\mathbf{A}\boldsymbol{x} - \langle \mathbf{A}\boldsymbol{x} \rangle)(\mathbf{A}\boldsymbol{x} - \langle \mathbf{A}\boldsymbol{x} \rangle)^{\mathrm{t}} \right\rangle$$
$$= \mathbf{A}\mathbf{C}_x\mathbf{A}^t. \qquad (14)$$

In this derivation $\mathbf{A}$ must be independent of the ensemble average. If we apply the above ansatz to the 2PCF, it seems reasonable to define a scaling relation for parameter dependent covariances as

$$\mathbf{C}_{\xi_{\pi}} = \mathbf{A}\mathbf{C}_{\xi_{\pi_0}}\mathbf{A}^{\mathrm{t}}, \qquad (15)$$
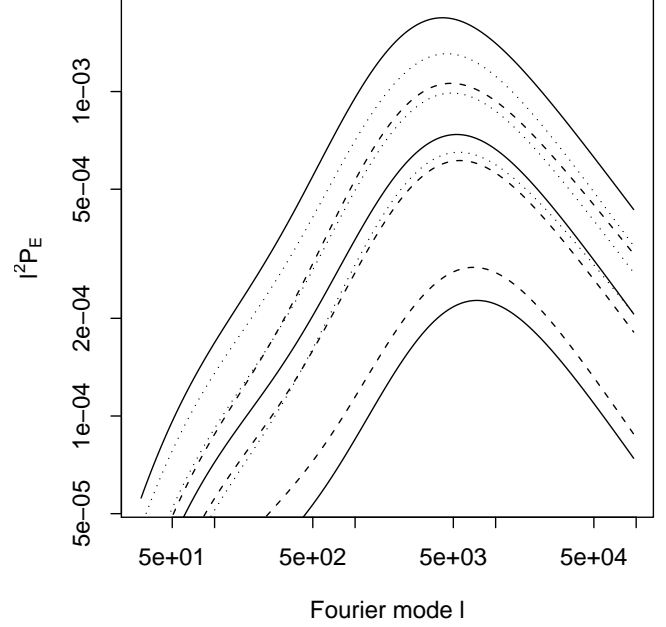


**Fig. 1.** The dimensionless shear power spectrum $\ell^2 P_{\mathrm{E}}$. The solid curves correspond to variation in $\Omega_{\mathrm{m}}$ and $\sigma_8$: $\Omega_{\mathrm{m}} = 0.2$, $\sigma_8 = 0.6$ (*lower*), $\Omega_{\mathrm{m}} = 0.3$, $\sigma_8 = 0.8$ (*middle*), $\Omega_{\mathrm{m}} = 0.4$, $\sigma_8 = 1.0$ (*top*). The dashed curves show variation in $\sigma_8$ with $\Omega_{\mathrm{m}} = 0.25$: $\sigma_8 = 0.6$ (*lower*), $\sigma_8 = 0.8$ (*middle*), $\sigma_8 = 1.0$ (*top*). The dotted curves show variation in $\Omega_{\mathrm{m}}$ with a constant $\sigma_8 = 0.9$: $\Omega_{\mathrm{m}} = 0.2$ (*lower*), $\Omega_{\mathrm{m}} = 0.3$ (*middle*), $\Omega_{\mathrm{m}} = 0.4$ (*top*).

where we can calculate the scaling matrices $\mathbf{A}$ using the 2PCF

$$diag(\mathbf{A}) = \boldsymbol{\xi}_{\pi}/\boldsymbol{\xi}_{\pi_0}. \qquad (16)$$

In contrast to a covariance matrix, the 2PCF can be calculated extremely fast for many different cosmologies via (3). Hence, it would be a fast and convenient method to calculate the covariance for a reference cosmology and then apply (15) to obtain covariances for arbitrary cosmological parameters. Unfortunately, we cannot transfer this method directly to the cosmic shear case. Recall, that the 2PCF is derived from the measured ellipticities of galaxies. Schneider et al. (2002b) have shown that the intrinsic ellipticity terms cancel out in the derivation of the 2PCF estimator, hence the 2PCF is defined only in terms of the shear. In contrast, the 2PCF covariance does not only consist of terms coming from the shear, but has additional noise terms which arise from the intrinsic ellipticity of galaxies. The pure shot noise term (8) is independent of cosmology and, as can be seen from (11-13), the mixed term cannot be scaled with relation (15), which is quadratic in the 2PCF.

However, in the limit of a noise-free covariance, i.e. considering only the cosmic variance term, a scaling relation similar to (15) exists. We explicitly prove this below, in particular, we show that the scaling matrices are independent of the ensemble average. The cosmic variance term can be calculated via (9). Cosmology only enters with the power spectrum, hence the relation of $\boldsymbol{\pi}$ to $\boldsymbol{\pi}_0$ can be described as $P_{\mathrm{E}}(\ell, \boldsymbol{\pi}) = a(\ell, \boldsymbol{\pi})P_{\mathrm{E}}(\ell, \boldsymbol{\pi}_0)$.

Using this relation we transform the cosmic variance term (9) for given bins $\vartheta_i, \vartheta_j$ as follows

$$
\begin{aligned}
V_{\pm\pm}(\boldsymbol{\pi}) &= \frac{1}{\pi A} \int_0^\infty \mathrm{d}\ell\, \ell\, \mathrm{J}_{0/4}(\ell\vartheta_i)\, \mathrm{J}_{0/4}(\ell\vartheta_j)\, P_{\mathrm{E}}^2(\ell, \boldsymbol{\pi}) \\
&= \frac{1}{\pi A} \sum_{\bar{\ell}} \Delta\bar{\ell}\, \bar{\ell}\, \mathrm{J}_{0/4}(\bar{\ell}\vartheta_i)\, \mathrm{J}_{0/4}(\bar{\ell}\vartheta_j)\, a^2(\bar{\ell}, \boldsymbol{\pi}) P_{\mathrm{E}}^2(\bar{\ell}, \boldsymbol{\pi}_0), \quad (17)
\end{aligned}
$$

where we discretize the integral into a sum of $\bar{\ell}$-bins. Next we insert equation (26) of Joachimi et al. (2008) (see also Kaiser 1998) but with $\sigma_\epsilon = 0$

$$
\left\langle \Delta P_{\mathrm{E}}(\bar{\ell}) \Delta P_{\mathrm{E}}(\bar{\ell}') \right\rangle = \frac{4\pi}{A\bar{\ell}\Delta\bar{\ell}} P_{\mathrm{E}}^2(\bar{\ell})\, \delta_{\bar{\ell}\bar{\ell}'}, \quad (18)
$$

to rewrite (17) as

$$
\begin{aligned}
V_{\pm\pm}(\boldsymbol{\pi}) &= \frac{1}{4\pi^2} \sum_{\bar{\ell}, \bar{\ell}'} \Delta\bar{\ell}^2\, \bar{\ell}'\, \bar{\ell}\, \mathrm{J}_{0/4}(\bar{\ell}\vartheta_i)\, \mathrm{J}_{0/4}(\bar{\ell}'\vartheta_j)\, a(\bar{\ell}, \boldsymbol{\pi})\, a(\bar{\ell}', \boldsymbol{\pi}) \\
&\quad \times \left\langle \Delta P_{\mathrm{E}}(\bar{\ell}, \boldsymbol{\pi}_0) \Delta P_{\mathrm{E}}(\bar{\ell}', \boldsymbol{\pi}_0) \right\rangle. \quad (19)
\end{aligned}
$$

The mean value theorem guarantees that there exist values $\bar{a}(\vartheta_i, \boldsymbol{\pi})$, $\bar{a}(\vartheta_j, \boldsymbol{\pi})$ such that (19) becomes

$$
\begin{aligned}
V_{\pm\pm}(\boldsymbol{\pi}) &= \frac{\bar{a}(\vartheta_i, \boldsymbol{\pi})\, \bar{a}(\vartheta_j, \boldsymbol{\pi})}{4\pi^2} \left\langle \int_0^\infty \mathrm{d}\ell\, \ell\, \mathrm{J}_{0/4}(\ell\vartheta_i)\, \Delta P_{\mathrm{E}}(\ell, \boldsymbol{\pi}_0) \right. \\
&\quad \times \left. \int_0^\infty \mathrm{d}\ell'\, \ell'\, \mathrm{J}_{0/4}(\ell'\vartheta_j)\, \Delta P_{\mathrm{E}}(\ell', \boldsymbol{\pi}_0) \right\rangle \\
&= \bar{a}(\vartheta_i, \boldsymbol{\pi})\, \bar{a}(\vartheta_j, \boldsymbol{\pi})\, V_{\pm\pm}(\boldsymbol{\pi}_0), \quad (20)
\end{aligned}
$$

where we consider the limit $\Delta\bar{\ell} \longrightarrow 0$ in the first step. Comparing the expressions of $V_{\pm\pm}(\boldsymbol{\pi})$ and $V_{\pm\pm}(\boldsymbol{\pi}_0)$ we can calculate the scaling factors as

$$
\bar{a}_{\pm\pm}(\vartheta_i, \boldsymbol{\pi}) = \frac{\int_0^\infty \mathrm{d}\ell\, \ell\, \mathrm{J}_{0/4}(\ell\vartheta_i) P_{\mathrm{E}}(\ell, \boldsymbol{\pi})}{\int_0^\infty \mathrm{d}\ell\, \ell\, \mathrm{J}_{0/4}(\ell\vartheta_i) P_{\mathrm{E}}(\ell, \boldsymbol{\pi}_0)} = \frac{\xi(\vartheta_i, \boldsymbol{\pi})}{\xi(\vartheta_i, \boldsymbol{\pi}_0)}, \quad (21)
$$

where we inserted (3) in the last step. This provides a fast and convenient method to scale the cosmic variance term in parameter space, due to the fact we can use a computationally efficient Hankel transformation for the calculation of the 2PCF. From (11) - (13) we see that the mixed term $M_{\pm\pm}$ scales linearly with the 2PCF which prevents a scaling relation similar to (15). Fortunately, the direct calculation of the linear term via (10) is comparatively fast, therefore, the scaling relation for the cosmic variance term already reduces the computational costs significantly.

Nonetheless, we numerically derive a fit-formula for the linear term based on the following expression

$$
M_{\pm\pm}(\vartheta_i, \vartheta_j, \boldsymbol{\pi}) = M_{\pm\pm}(\vartheta_i, \vartheta_j, \boldsymbol{\pi}_0) \left(\frac{\Omega_{\mathrm{m}}}{0.25}\right)^\alpha \left(\frac{\sigma_8}{0.9}\right)^\beta. \quad (22)
$$

The structure of this fit-formula is motivated by the intention to use as few fit-parameters as possible; additionally we require that in the limit of the fiducial model, $M_{\pm\pm}(\vartheta_i, \vartheta_j, \boldsymbol{\pi}) = M_{\pm\pm}(\vartheta_i, \vartheta_j, \boldsymbol{\pi}_0)$ must hold. The fit-parameters $\alpha$ and $\beta$ vary depending on the scale $\vartheta_i, \vartheta_j$ and are different for the different parts of the covariance matrix, $\mathbf{C}_{++}$, $\mathbf{C}_{--}$, and $\mathbf{C}_{+-}$. The tables with $\alpha$ and $\beta$ are available on the internet[1].

---
[1] http://www.astro.uni-bonn.de/~teifler/fit-parameters.pdf

## 3.2. Variation of the inverse covariance with $\Omega_{\mathrm{m}}$ and $\sigma_8$

From the variation of the power spectrum with $\Omega_{\mathrm{m}}$ and $\sigma_8$ (Sect. 3.1) it is clear that covariances vary with respect to comological parameters. For simplicity and in order to increase the readability of the following sections we refer to this variation as CDC-effect (CDC ≡ Cosmology Dependent Covariances). In order to examine the CDC-effect more closely, recall that the structure of the covariance is given by

$$
\mathbf{C} = \left( \begin{array}{c|c} \mathbf{C}_{++} & \mathbf{C}_{+-} \\ \hline \mathbf{C}_{+-}^{\mathrm{t}} & \mathbf{C}_{--} \end{array} \right) \quad (23)
$$

and the individual parts are calculated from (5) - (10). From these equations we see that the covariances are filtered versions of the power spectrum, either filtered by a product of $\mathrm{J}_0$'s (in case of $\mathbf{C}_{++}$), $\mathrm{J}_4$'s ($\mathbf{C}_{--}$), or a combination of both ($\mathbf{C}_{+-}$). The strength of the CDC-effect depends on these filter functions, as they determine which parts of the power spectrum are sampled. A change in $\Omega_{\mathrm{m}}$ and $\sigma_8$ affects all scales of the power spectrum almost similarly (see Fig. 1) therefore, the CDC-effect for the individual parts of $\mathbf{C}$ is also similar. However, this might change when considering different cosmological parameters, e.g. the shape parameter $\Gamma$. A change in $\Gamma$ rotates the power spectrum. The covariances are integrals over $P_{\mathrm{E}}$, and depending on the filter function, the change in $P_{\mathrm{E}}$ can average out. A second argument why the individual covariance parts have different sensitivity to the CDC-effect is that $\mathbf{C}_{+-}$ is not affected from shot noise, hence a change in cosmology has a stronger impact on $\mathbf{C}_{+-}$ compared to $\mathbf{C}_{++}$ and $\mathbf{C}_{--}$.

In order to quantify the CDC-effect we examine the trace of the inverse covariance matrix $\mathbf{C}^{-1}$. The trace of the covariance itself is an improper measure for this effect, as it depends on the binning, which can be seen from (8). The trace of $\mathbf{C}$ becomes arbitrarily large when decreasing the bin width. In contrast, we checked numerically that for the trace of $\mathbf{C}^{-1}$ binning effects are negligible, once one has exceeded a minimum bin number. More precisely, once the bin width of the 2PCF data vector is small enough that discretization effects are unimportant, the trace of $\mathbf{C}^{-1}$ hardly changes for different binning.

Figure 2 shows the trace of the inverse covariance matrix depending on $\Omega_{\mathrm{m}}$ for various constant $\sigma_8$ (top) and vice versa (bottom). Here, we normalize the survey size to $A = 1\,\mathrm{deg}^2$; the other survey parameters are $\sigma_\epsilon = 0.4$ and $\bar{n} = 10/\mathrm{arcmin}^2$. We postpone a detailed analysis of how survey parameters influence the CDC-effect to Sect. 4. Qualitatively the result does not change for different survey parameters; the trace of $\mathbf{C}^{-1}$ decreases with increasing $\Omega_{\mathrm{m}}$ or $\sigma_8$.

In addition, we perform a singular value decomposition (SVD) for each inverse covariance matrix. For the case of a symmetric and positive definite matrix, such as the inverse covariance matrix, an SVD yields the eigenvalues in decreasing order. For arbitrary $i$, we find that the $i$-th eigenvalue decreases when increasing $\Omega_{\mathrm{m}}$ or $\sigma_8$. The strength of the CDC-effect, i.e. the gradient of the traces, depends on the considered point in parameter space.
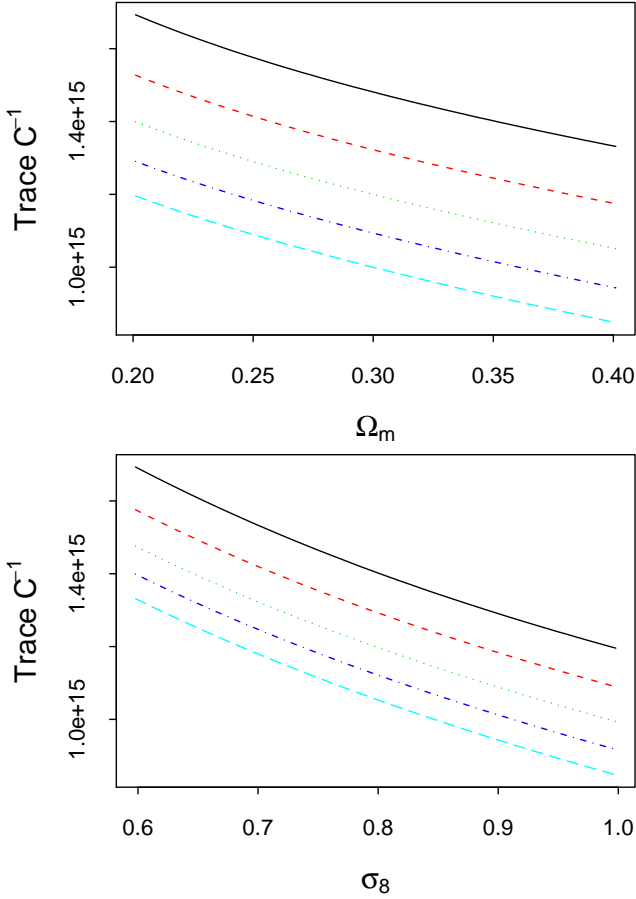
**Fig. 2.** The trace of the inverse covariance matrix $\mathbf{C}^{-1}$ depending on $\Omega_m$ (*top*), the individual lines in each figure correspond to (from top to bottom) $\sigma_8 = [0.6, 0.7, 0.8, 0.9, 1.0]$. The *lower* panel shows the dependence on $\sigma_8$, the individual lines corresponding to (from top to bottom) $\Omega_m = [0.2, 0.25, 0.3, 0.35, 0.4]$.

# 4. Impact of the CDC-effect on parameter estimation

## 4.1. Basics of the likelihood analysis

Throughout the whole likelihood analysis we assume the $\Lambda$CDM model. We define the posterior likelihood $p(\boldsymbol{\pi}|\boldsymbol{\xi})$ for the case of a 2PCF data vector as

$$p(\boldsymbol{\pi}|\boldsymbol{\xi}) = \frac{p(\boldsymbol{\xi}|\boldsymbol{\pi})}{p(\boldsymbol{\xi})}\, p(\boldsymbol{\pi})\,, \qquad (24)$$

where $p(\boldsymbol{\pi})$ denotes the prior probability density, $p(\boldsymbol{\xi}|\boldsymbol{\pi})$ is the likelihood and $p(\boldsymbol{\xi})$ denotes the evidence. The prior usually contains knowledge on the parameter vector $\boldsymbol{\pi}$ coming from former experiments. Here, we assume flat priors with cutoffs, which means $p(\boldsymbol{\pi})$ is constant for all parameters inside a fixed interval (i.e. $\Omega_m \in [0.2; 0.4]$, $\sigma_8 \in [0.6; 1.0]$) and $p(\boldsymbol{\pi}) = 0$ else. We assume that $\boldsymbol{\xi}$ is normally distributed in parameter space, hence our likelihood $p(\boldsymbol{\xi}|\boldsymbol{\pi})$ can be written as

$$p(\boldsymbol{\xi}|\boldsymbol{\pi}) = \frac{\exp\left[-\frac{1}{2}\left((\boldsymbol{\xi}_{\boldsymbol{\pi}} - \hat{\boldsymbol{\xi}})^t\,\mathbf{C}^{-1}\,(\boldsymbol{\xi}_{\boldsymbol{\pi}} - \hat{\boldsymbol{\xi}})\right)\right]}{(2\pi)^{d/2}\,|\mathbf{C}|^{\frac{1}{2}}}\,, \qquad (25)$$

where $\hat{\boldsymbol{\xi}}$ denotes the mean data vector, $\boldsymbol{\xi}_{\boldsymbol{\pi}}$ the model data vector, $d$ is the dimension of the data vectors, hence $|\mathbf{C}|$ is the determinant of a $d{\times}d$ covariance matrix. Note that the 2PCF data vector consists of two parts ($\boldsymbol{\xi}_+$ and $\boldsymbol{\xi}_-$), each with $d/2$ bins. The evidence is a normalization obtained by integrating the likelihood over the considered parameter space

$$p(\boldsymbol{\xi}) = \int d\boldsymbol{\pi}\,\frac{\exp\left[-\frac{1}{2}\left((\boldsymbol{\xi}_{\boldsymbol{\pi}} - \hat{\boldsymbol{\xi}})^t\,\mathbf{C}^{-1}\,(\boldsymbol{\xi}_{\boldsymbol{\pi}} - \hat{\boldsymbol{\xi}})\right)\right]}{(2\pi)^{d/2}\,|\mathbf{C}|^{\frac{1}{2}}}\,. \qquad (26)$$

In our case we calculate $\hat{\boldsymbol{\xi}}$ from $P_E$ via (3) assuming our fiducial cosmology; $\boldsymbol{\xi}_{\boldsymbol{\pi}}$ is calculated similarly but its cosmological model varies according to the considered point in parameter space. The result of a likelihood analysis is usually summarized in contour plots. In a Bayesian approach, these likelihood contours represent so-called credible regions, i.e. a region in parameter space, where the true parameter is located with a probability of 68%, 95%, 99,9%, respectively. In addition, we quantify the size of these credible regions through the determinant of the second-order moment of the posterior likelihood (see Kilbinger & Schneider 2004)

$$Q_{ij} \equiv \int d^2\boldsymbol{\pi}\, p(\boldsymbol{\pi}|\boldsymbol{\xi})\, (\pi_i - \pi_i^f)(\pi_j - \pi_j^f), \qquad (27)$$

with $\pi_1$ and $\pi_2$ as the varied parameters, $\pi_i^f$ as the parameter of the fiducial model. The determinant is given by

$$q = \sqrt{|Q_{ij}|} = \sqrt{Q_{11}Q_{22} - Q_{12}^2}. \qquad (28)$$

Smaller credible regions in parameter space correspond to a smaller value of $q$. In this paper all $q$ are given in units of $10^{-4}$.

## 4.2. Results of the likelihood analysis

In Sect. 3 we calculate 2500 covariances covering a parameter range of $\Omega_m \in [0.2; 0.4]$ and $\sigma_8 \in [0.6; 1.0]$. Here, we want to examine how the CDC-effect influences the likelihood contours, hence for each of the 2500 covariance matrices we perform a likelihood analysis. In these analyses, we consider the same parameter space, similar priors, similar $\hat{\boldsymbol{\xi}}$ and $\boldsymbol{\xi}_{\boldsymbol{\pi}}$, only the covariance in (25) is changed. The left panel of Fig. 3 shows the 95%-credible intervals when choosing $\Omega_m = 0.2$ and $\sigma_8 = 0.6$ (solid), $\Omega_m = 0.4$, and $\sigma_8 = 1.0$ (dotted) as a model for the covariance matrix. We compare these to the (dashed) case when the covariance is calculated from the fiducial model ($\Omega_m = 0.25$, $\sigma_8 = 0.9$). These examples illustrate that assuming different cosmologies in the covariance can significantly broaden or narrow the likelihood contours. As expected from the foregoing analysis of the inverse covariance traces (Sect. 3) the contours broaden for increasing $\Omega_m$ and $\sigma_8$.

Without any information which cosmology to choose in our covariance matrix, it is reasonable to include prior information coming from other cosmological probes into our covariance cosmology. The middle panel of Fig. 3 shows the 95% credible intervals when calculating the covariance from the minimum, mean, and maximum values of the 68% confidence region of the recent WMAP 5-years analysis (Komatsu et al. 2008). Compared to the left panel the deviation of the contours
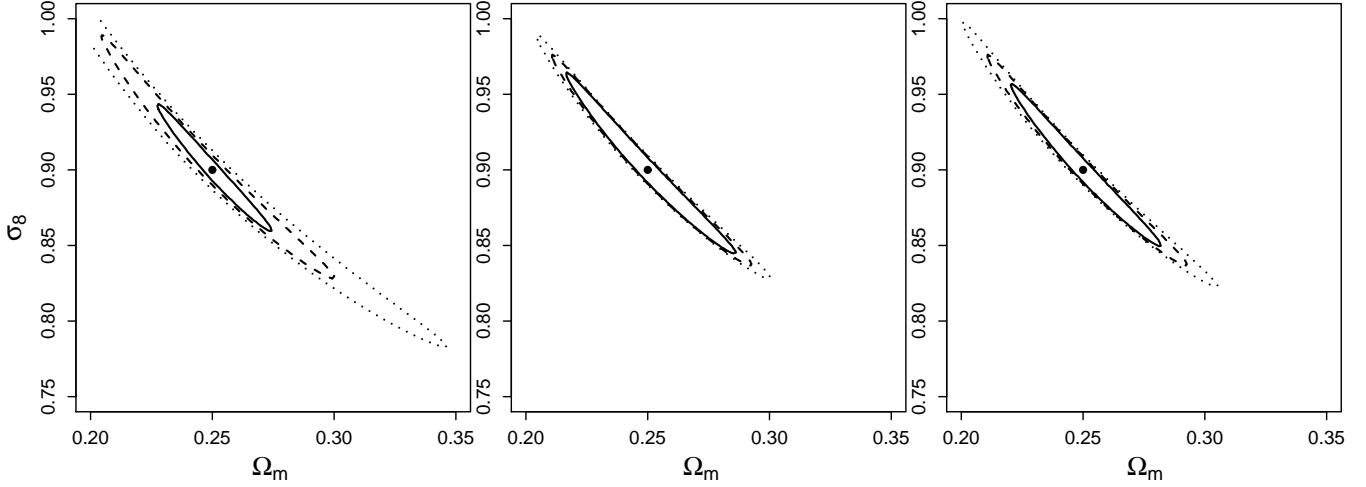
**Fig. 3.** The 95%-credible intervals obtained from likelihood analyses with different cosmological models assumed in their covariance matrix. The left panel corresponds to the following covariance parameters: $\Omega_m = 0.2$, $\sigma_8 = 0.6$ (*solid*), $\Omega_m = 0.25$, $\sigma_8 = 0.9$ (*dashed*), and $\Omega_m = 0.4$, $\sigma_8 = 1.0$ (*dotted*). The middle panel shows the deviation which occurs when restricting the range of possible covariance models to the 68% confidence interval of the WMAP5 analysis, i.e. $\Omega_m = 0.237$, $\sigma_8 = 0.74$ (*solid*), $\Omega_m = 0.259$, $\sigma_8 = 0.796$ (*dashed*), and $\Omega_m = 0.274$, $\sigma_8 = 0.85$ (*dotted*). The right panel shows the same analysis but for the 95% confidence interval of the WMAP5 analysis, i.e. $\Omega_m = 0.226$, $\sigma_8 = 0.70$ (*solid*), $\Omega_m = 0.237$, $\sigma_8 = 0.74$ (*dashed*), and $\Omega_m = 0.288$, $\sigma_8 = 0.885$ (*dotted*).

**Table 1.** Values of $q$ for different covariance models

| parameters used for the covariance | $q$ $[10^{-4}]$ |
|---|---|
| $\Omega_m = 0.25$, $\sigma_8 = 0.9$ | 1.03 |
| $\Omega_m = 0.2$, $\sigma_8 = 0.6$ | 0.28 |
| $\Omega_m = 0.4$, $\sigma_8 = 1.0$ | 3.30 |
| $\Omega_m = 0.259$, $\sigma_8 = 0.796$ (WMAP5 68 % CL mean) | 0.75 |
| $\Omega_m = 0.237$, $\sigma_8 = 0.740$ (WMAP5 68 % CL min) | 0.56 |
| $\Omega_m = 0.274$, $\sigma_8 = 0.850$ (WMAP5 68 % CL max) | 1.02 |
| $\Omega_m = 0.226$, $\sigma_8 = 0.700$ (WMAP5 95 % CL min) | 0.45 |
| $\Omega_m = 0.288$, $\sigma_8 = 0.885$ (WMAP5 95 % CL max) | 1.24 |

reduces significantly, nevertheless it is still noticeable and cannot be neglected in a precision cosmology analysis. Similarly, the right panel shows the impact of the CDC-effect when calculating the covariance from parameters within the 95% confidence region of the recent WMAP5 analysis. For a better comparison we calculate the values of $q$ (Sect. 4.1) for all contour plots and summarize them in Table 1. Restricting the possible cosmologies for the covariance to the 68% contour region of the WMAP5 analysis, the values of $q$ deviate by a factor of $\approx 1.84$. This factor increases to $\approx 2.76$ when considering the minimum and maximum values of the 95% confidence region of the WMAP5 constraints. In Fig. 4 we show the values of $q$ for all 2500 likelihood analyses depending on $\Omega_m$ (top) and $\sigma_8$ (bottom). Similar to the parameter dependence of the inverse covariances in Sect. 3, the strength of the CDC-effect, i.e. the gradient of the curves in Fig. 4, depends on the considered point in parameter space. At the fiducial model we calculate $(\partial q/\partial\Omega_m)_{fid} = 7.5$, whereas in case of $\sigma_8$ we find $(\partial q/\partial\sigma_8)_{fid} = 3.5$.

### 4.3. Impact of survey parameters on the CDC-effect

In the last section we have shown, that the CDC-effect non-negligibly affects the likelihood contours. However, we only quantify this for one specific set of survey parameters. In this section we examine how the impact of the CDC-effect on likelihood contours depends on survey parameters, namely survey size $A$, ellipticity dispersion $\sigma_\epsilon$, and number density of source galaxies $\bar{n}$, where in case of the latter two only the combination $\sigma_\epsilon^2/\bar{n}$ is of interest. We perform likelihood analyses for 9 different combinations of $\sigma_\epsilon^2/\bar{n}$ and 8 different survey sizes. The strength of the CDC-effect is quantified by the ratio of maximum to minimum value of $q$, which occur within the considered range of $\Omega_m$ and $\sigma_8$, we define $\Delta q = q(\pi_{max})/q(\pi_{min})$. The minimum $q$ is obtained when choosing the minimum parameter set in the calculation of the covariance, i.e. $\pi_{min} = (\Omega_m = 0.2, \sigma_8 = 0.6)$. Correspondingly, choosing the maximum parameter set $\pi_{max} = (\Omega_m = 0.4, \sigma_8 = 1.0)$ results in the maximal $q$. The values of $q$ represent the size of credible intervals, hence $\Delta q$ can be interpreted as their ratio.

Unfortunately, it is not possible to derive an analytical expression for the relation between $\Delta q$ and the survey parameters. From (8) - (10) we see that the individual covariance terms scale differently with $\sigma_\epsilon^2/\bar{n}$. This already prohibits an analytically derived relation between $\Delta q$ and $\sigma_\epsilon^2/\bar{n}$. Considering the survey size $A$, (8) - (10) imply that the total covariance scales with $1/A$. When comparing two (inverse) covariances with different cosmologies by taking their ratio, the survey size cancels, suggesting the strength of CDC-effect to be independent of $A$. However, when considering the likelihood, the inverse covariance enters in the exponent, furthermore the values of $q$ are an integral over the posterior likelihood. This non-linearity in the inverse covariance causes that the strength of the CDC-
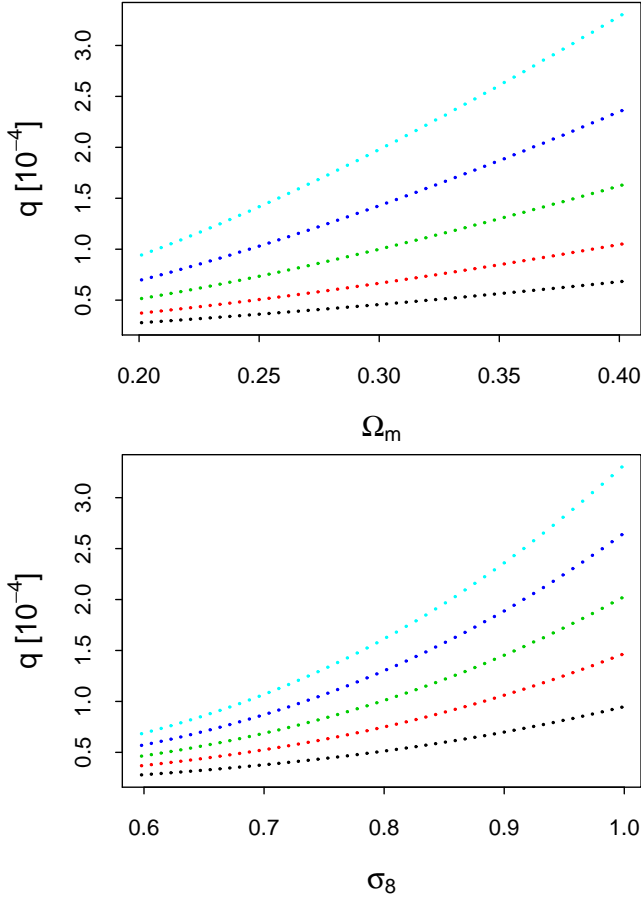
**Fig. 4.** The values of $q$ depending on $\Omega_m$ (*top*), the individual lines in each figure correspond to (from top to bottom) $\sigma_8 = [0.6, 0.7, 0.8, 0.9, 1.0]$. The *lower* panel shows the dependence on $\sigma_8$, the individual lines corresponding to (from top to bottom) $\Omega_m = [0.2, 0.25, 0.3, 0.35, 0.4]$.
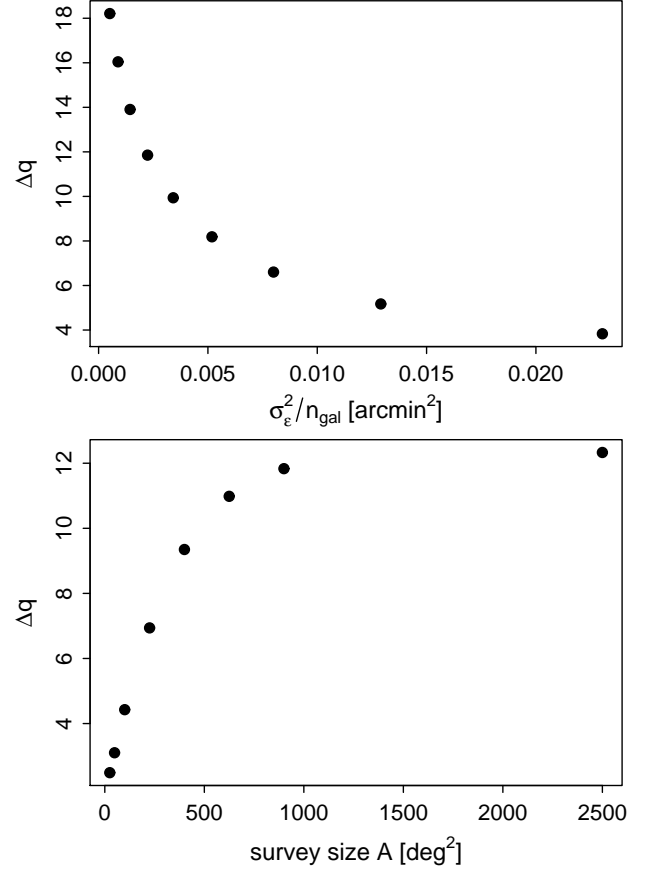
**Fig. 5.** The ratio of maximum to minimum value of $q$ depending on the ratio $\sigma_\epsilon^2/\bar{n}$ (*upper panel*) and depending on the survey size $A$ (*lower panel*).

effect varies with the survey size. An analytic expression of this dependence cannot be derived, for similar reasons as for the case of $\sigma_\epsilon^2/\bar{n}$. We therefore calculate $\Delta q$ depending on the survey parameters numerically.

The upper panel of Fig. 5 shows $\Delta q = q(\pi_{max})/q(\pi_{min})$ as a function of $\sigma_\epsilon^2/\bar{n}$. The ratio $\Delta q$ changes from 4 to 18 over the considered interval of $\sigma_\epsilon^2/\bar{n}$. When increasing the survey size $A$ (Fig. 5, lower panel), we find that the impact of the CDC-effect increases from $\Delta q = 2.5$ (for a 25 deg$^2$ survey) up to $\Delta q = 12.3$ (for a 2500 deg$^2$ survey). Note that the size of the likelihood contours, hence the values of $q$ themselves, decrease with decreasing $\sigma_\epsilon^2/\bar{n}$ and increasing $A$. In contrast, $\Delta q$ increases with decreasing $\sigma_\epsilon^2/\bar{n}$ and increasing $A$. Hence relatively, the CDC-effect becomes more important when increasing the survey size or when decreasing the ratio $\sigma_\epsilon^2/\bar{n}$.

## 5. Likelihood analysis with a model dependent covariance

### 5.1. Adaptive covariance matrix

For a given cosmological model we can calculate the covariance directly from (5) - (10). This enables us to perform a likelihood analysis, where the covariance is calculated individually for every point in parameter space. We denote this parameter dependent covariance as $\mathbf{C}_\pi$ and rewrite the likelihood (25) as

$$p(\xi|\pi) = \frac{\exp\left[-\frac{1}{2}\left((\xi_\pi - \hat{\xi})^t \, \mathbf{C}_\pi^{-1} \, (\xi_\pi - \hat{\xi})\right)\right]}{(2\pi)^{d/2} \, |\mathbf{C}_\pi|^{\frac{1}{2}}} . \tag{29}$$

Compared to the case of a constant covariance, there are two main differences. First, the covariance in the exponential term of (29) changes according to the considered point in parameter space. Second, $|\mathbf{C}_\pi|^{\frac{1}{2}}$ is now parameter dependent, therefore the determinant no longer cancels with a similar term in the evidence. As a consequence, the posterior likelihood does not only depend on the exponential terms, which basically compare $\xi_\pi$ and $\hat{\xi}$, but it is also affected by the determinants of the covariance matrices, more precisely by their behavior in parameter space. In the following we quantify the impact of the determinant term.

The upper left panel in Fig. 6 shows the likelihood contours for a 84 deg$^2$ survey, where the posterior probability is calculated via the new likelihood (29). For comparison, the right panel shows the likelihood contours when neglecting the parameter dependence in the determinant terms, hence considering a parameter dependent covariance only in the exponential terms. One clearly sees that the determinant terms shift the likelihood contours and cause a difference between the best-fit value and the fiducial model. In order to explain this shift we overlay the right panels of Fig. 6 with the contours of constant $|\mathbf{C}_\pi|^{-1/2}$ (for numerical reasons we plot $\ln |\mathbf{C}_\pi|^{-1/2}$). We see, that the covariance determinant is a monotonic function of $\Omega_m$ and $\sigma_8$; it decreases with increasing $\Omega_m$ or $\sigma_8$. Hence, $|\mathbf{C}_\pi|^{-\frac{1}{2}}$ induces a parameter-dependent weighting, which increases the likelihood at small $\Omega_m$ and $\sigma_8$ and vice versa suppresses large $\Omega_m$ and $\sigma_8$. In general, the exponential term dominates the likelihood, $|\mathbf{C}_\pi|^{-1/2}$ only has significant impact on parameter regions where the exponential hardly changes. For the highly degenerate case of $\Omega_m$ and $\sigma_8$, this applies to curves where $\sigma_8 \approx const \times \Omega_m^{-0.6}$. Compared to these curves, the contours of constant $|\mathbf{C}_\pi|^{-1/2}$ are slightly rotated, which allows for different values of the latter in regions where the exponential term is constant. As a result, the likelihood contours in the left panel are shifted and stretched towards regions of larger $|\mathbf{C}_\pi|^{-1/2}$ compared to the right panel. Note that for a different parameter combination this bias might not cause such a large shift of the best-fit value.

The second row of Fig. 6 shows the same analysis but for a 900 deg$^2$ survey. Comparing the left and right panel, we see that the likelihood contours are, similar to the 84 deg$^2$ survey, shifted and stretched towards regions of larger $|\mathbf{C}_\pi|^{-1/2}$. However, the effect is hardly noticeable and the bias of the best-fit value has basically vanished. This can be explained when looking at the expression of the posterior likelihood

$$p(\pi|\xi) = \frac{\exp\left[-\frac{1}{2}\left((\xi_\pi - \hat{\xi})^t \, \mathbf{C}_\pi^{-1} \, (\xi_\pi - \hat{\xi})\right)\right]}{\int d\pi' |\mathbf{C}_{\pi'}^{-1} \, \mathbf{C}_\pi|^{\frac{1}{2}} \exp\left[-\frac{1}{2}\left((\xi_{\pi'} - \hat{\xi})^t \, \mathbf{C}_{\pi'}^{-1} \, (\xi_{\pi'} - \hat{\xi})\right)\right]}. \quad (30)$$

Compared to the case of a constant covariance the above expression has an additional factor in the denominator, i.e. $|\mathbf{C}_\pi \mathbf{C}_{\pi'}^{-1}|^{1/2}$. Note, that this factor is independent of the survey size $A$, whereas the importance of the exponential term increases with increasing $A$. As a result, the cosmology dependence of the covariance determinant becomes negligible for sufficiently large surveys.

### 5.2. Fisher matrix analysis

We expect tighter constraints on cosmological parameters if the cosmology dependence of both, mean data vector and covariance matrix, is incorporated into the likelihood analysis, instead of only using the mean data vector (Tegmark et al. 1997). The Fisher information matrix can be used to illustrate this effect; its definition reads (Kendall & Stuart 1979; Tegmark et al. 1997)

$$\mathbf{F}_{ij} = \left\langle \frac{\partial^2 L(\pi)}{\partial \pi_i \, \partial \pi_j} \right\rangle = \left( \frac{\partial^2 L(\pi)}{\partial \pi_i \, \partial \pi_j} \right)_{\pi = \pi_{ML}}, \quad (31)$$

where $L = -\ln p(\pi|\xi)$, $\pi = (\pi_1, ..., \pi_n)$ describes the underlying (cosmological) parameters, and $\pi_{ML}$ denotes the maximum likelihood parameter vector.

In addition, if we Taylor-expand $L(\pi)$ in parameter space at the fiducial parameters we derive

$$L(\pi) = L(\pi_{fid}) + 0 + \frac{1}{2}(\pi - \pi_{fid})^t \, \mathbf{T}^{-1}(\pi - \pi_{fid}) + O(\Delta_\pi^3), \quad (32)$$

with

$$(\mathbf{T}^{-1})_{ij} = \left( \frac{\partial^2 L(\pi)}{\partial \pi_i \, \partial \pi_j} \right)_{\pi = \pi_{fid}}. \quad (33)$$

The first-order term vanishes since $(\partial L/\partial \pi)|_{\pi_{fid}}$ is zero, hence (32) is dominated by second-order terms. In this analysis we only consider $\Omega_m$ and $\sigma_8$; $\pi_{ML}$ corresponds to our fiducial model, i.e. $\pi_{fid} = (\Omega_m = 0.25, \sigma_8 = 0.9)$. Comparing (33) and (31) one sees that the Fisher matrix and the inverse parameter covariance matrix $\mathbf{T}^{-1}$ are equal. We rewrite (32) as

$$L(\pi) = L(\pi_{fid}) + \frac{1}{2}(\pi - \pi_{fid})^t \, \mathbf{F} \, (\pi - \pi_{fid}) + O(\Delta_\pi^3). \quad (34)$$

This equation enables us, for a given Fisher matrix, to calculate lower bounds for $L(\pi)$, hence we can derive lower bounds on the likelihood contours. For the case that $p(\xi|\pi)$ is Gaussian, which at least close to the maximum likelihood parameter vector is a good approximation ($p(\xi|\pi) \propto \exp[-L]$), one can directly express the Fisher matrix in terms of the mean data vector and the data covariance matrix (e.g. Tegmark et al. 1997)

$$\mathbf{F}_{ij} = \frac{1}{2} tr \left[ \mathbf{C}^{-1} \mathbf{C}_{,i} \mathbf{C}^{-1} \mathbf{C}_{,j} + \mathbf{C}^{-1} \mathbf{M}_{ij} \right], \quad (35)$$

where $\mathbf{C}_{,i} \equiv \partial \mathbf{C}/\partial \pi_i$ denotes the derivative of the covariance matrix with respect to the $i$-th component of the parameter vector and $\mathbf{M}_{ij} \equiv \xi_{,i} \xi_{,j}^t + \xi_{,j} \xi_{,i}^t$. The first term of (35) vanishes in case the covariance matrix is constant in parameter space, the second term vanishes in case of a constant mean. For cosmic shear we have seen that neither the mean data vector, nor the covariance matrix are independent of cosmological parameters, hence when calculating the Fisher matrix both terms must be taken into account. Recall that $\mathbf{C} \propto 1/A$, which also holds for the derivatives $\mathbf{C}_{,i}$, hence the first term is independent of the survey size. The second term increases proportional to the survey volume, therefore the information gain on cosmological parameters, through incorporating the cosmology dependence of covariances, becomes less important for large surveys.

Figure 7 shows the results of the Fisher matrix analysis for two different survey sizes (84 deg$^2$ on the left and 900 deg$^2$ on the right). As expected, the left panel (smaller survey) shows a small improvement, which vanishes completely in case of the larger survey (right panel). Nevertheless, one should keep in mind that we only consider Gaussian covariances. The cosmology dependence of the covariance becomes larger for the case of non-Gaussian covariances for the following reason. Non-Gaussianity increases the cosmic variance term, in particular it becomes important on small scales, which are still dominated by shot noise in the pure Gaussian case. As the CDC-effect mainly results from the cosmic variance term, its strength also increases in the non-Gaussian case. A stronger dependence of
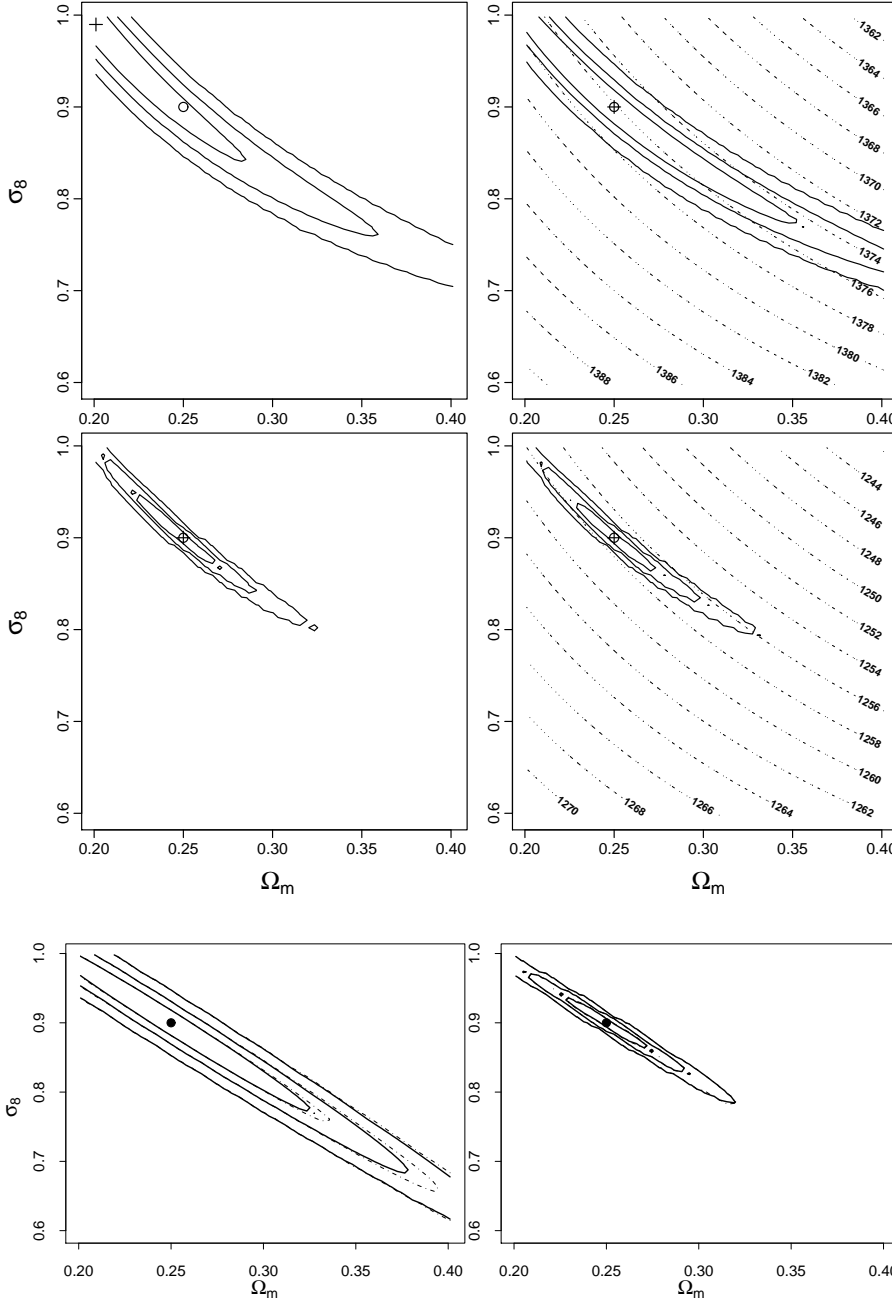
**Fig. 6.** The left plots shows the likelihood contours when using a model- dependent covariance, more explicitly, when calculating the posterior from (29). The cross illustrates the best-fit value, whereas the circle indicates our fiducial model. The panels on the r.h.s. show the likelihood contours obtained when neglecting the determinant-terms (29). The dotted contours visualize regions of constant $\ln |\mathbf{C}_{\pi'}|^{-1/2}$. The likelihood contours in the upper row correspond to a survey size of 84 $\deg^2$, whereas the lower panels correspond to $A = 900$ $\deg^2$.



**Fig. 7.** Likelihood contours from a Fisher matrix analysis for a 84 $\deg^2$ survey (*left*), and for a 900 $\deg^2$ (*right*). The dashed lines correspond to the same analysis but neglecting the covariance term. The dot indicates the fiducial model at which the Fisher matrix was calculated. Note, that in the right panel dashed and solid contours are identical

the covariance on parameters enlarges the first term in (35), which implies that for the case of truly non-Gaussian covariances the improvement on parameter constraints is more significant than shown in Fig. 7.

### 5.3. Iterative likelihood analysis

In Sect. 5.1 we have introduced the adaptive covariance, which is a proper way to incorporate cosmology dependent covariances into a likelihood analysis. Its disadvantage is the large computational effort, which is high already for Gaussian covariances. In order to account for non-Gaussianity, one must employ ray-tracing covariances derived from many numerical simulations with different underlying cosmologies. In a multi-dimensional parameter space, this is clearly unfeasible with to-

day's computer power.

In this section we quantify the impact on likelihood contours when using non-Gaussian instead of Gaussian covariances. We use a ray-tracing covariance taken from the Millennium simulation (Hilbert et al. 2008), neglect the CDC-effect and approximate the covariance to be constant in parameter space. The error in the posterior likelihood caused by this approximation increases with increasing distance to the cosmology of the ray-tracing simulation. As we are mainly interested in regions around the maximum likelihood parameter set, $\boldsymbol{\pi}_{\mathrm{ML}}$, this suggests the following strategy for a likelihood analysis. First, perform an iterative likelihood analysis using Gaussian covariances in order to derive $\boldsymbol{\pi}_{\mathrm{ML}}$. Then, start a numerical simulation with this cosmology, derive a ray-tracing covariance, and perform the final likelihood analysis. This ansatz minimizes the

**Table 2.** The ML-parameter sets which occur when choosing different starting cosmologies in the iterative likelihood analysis.

| step | run1 | | run2 | | run3 | | run4 | | run5 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\Omega_m$ | $\sigma_8$ | $\Omega_m$ | $\sigma_8$ | $\Omega_m$ | $\sigma_8$ | $\Omega_m$ | $\sigma_8$ | $\Omega_m$ | $\sigma_8$ |
| $\pi_{start}$ | 0.20 | 0.60 | 0.237 | 0.740 | 0.250 | 0.90 | 0.274 | 0.850 | 0.40 | 1.0 |
| $\pi_{ML}$ 1 | 0.254 | 0.892 | 0.254 | 0.892 | 0.245 | 0.914 | 0.259 | 0.884 | 0.277 | 0.858 |
| $\pi_{ML}$ 2 | 0.260 | 0.882 | 0.260 | 0.882 | 0.245 | 0.914 | 0.260 | 0.882 | 0.259 | 0.884 |
| $\pi_{ML}$ 3 | 0.260 | 0.882 | 0.260 | 0.882 | converged | | 0.260 | 0.882 | 0.260 | 0.882 |
| $\pi_{ML}$ 4 | converged | | converged | | converged | | converged | | 0.260 | 0.882 |
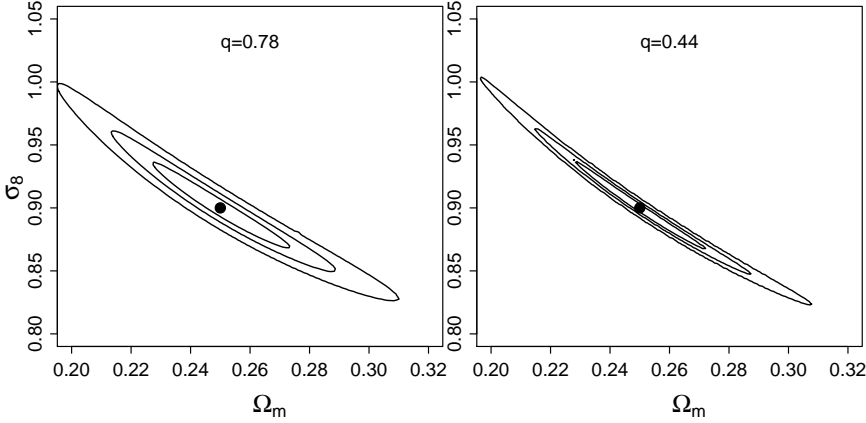


**Fig. 8.** The likelihood contours when using a ray-tracing covariance derived from the Millennium Simulation via field-to-field variation (*left panel*), compared to the case of a Gaussian covariance (*right panel*). Although the original size of each field is only 16 deg$^2$, we extrapolated the covariance to a 900 deg$^2$ survey. The values of $q$ are given in units of $10^{-4}$.

errors due to the CDC-effect in the region of interest and additionally incorporates non-Gaussianity.

In order to derive $\pi_{ML}$ iteratively, we start from an arbitrary cosmology, calculate a Gaussian covariance matrix therefrom using (5) - (10), and perform a likelihood analysis. Throughout this first iteration step the covariance matrix is kept constant. In the second step we choose the ML-parameter set of the first analysis as the underlying cosmology for the new covariance matrix, and again perform a likelihood analysis. We continue this iteration process until the ML-parameter set converges. The main difficulty of this ansatz is that the choice of the starting cosmology might influence the final ML-parameter estimate and therefore also the final covariance. In order to check for this, we take the noise of a ray-tracing data vector, add it to our fiducial data vector and thereby simulate measurement uncertainties in the latter. When performing the analysis without noise the iteration converges after one step, as the model data vector (at the fiducial model) exactly fits the fiducial data vector, $\xi_{\pi_{fid}} = \hat{\xi}$. Table 2 shows the results for 5 iterative likelihood analyses, each starting from a different cosmology in the covariance. We see that all 5 runs converge quickly, 4 of them to the same cosmology. Only the run which started from the fiducial model deviates from the others. Although the suggested $\pi_{ML}$ are close to $\pi_{fid}$, we note that none of the runs converges to the fiducial model. This implies that the starting cosmology can bias the final outcome of the iterative likelihood analysis and can shift the ML-estimate. In general, such a bias occurs if the function $\xi_\pi - \hat{\xi}$ does not fall off steeply enough around the ML-parameter set, which especially applies to higher-dimensional likelihood analyses.

Our iterative pre-analysis has converged to $\Omega_m = 0.26$, $\sigma_8 = $

0.882, however we "only" have a ray-tracing simulation with $\Omega_m = 0.25$, $\sigma_8 = 0.9$ available. Figure 8 shows the result of our likelihood analysis, when using the ray-tracing covariance of the Millennium simulation (left panel). Compared to a likelihood analysis using a Gaussian covariance (right panel), the contours broaden significantly; $q$ increases from $0.44 \times 10^{-4}$ in the Gaussian to $0.78 \times 10^{-4}$ in the non-Gaussian case. Note that the value of $q$ in the Gaussian case does not correspond to that in Table 1, because we use different survey parameters (here, $\sigma_\epsilon = 0.3$, $\bar{n} = 15/\text{arcmin}^2$) and a different data vector (here, 30 logarithmic bins from 0.2-130 arcmin) in order to exactly match the corresponding parameters of the ray-tracing covariance.

The impact of non-Gaussianity depends on the scales probed by the data vector. In our case 20 bins are below 10 arcmin, therefore the impact is relatively high. Choosing linear bins or probing higher $\vartheta$ reduces the difference to the Gaussian case. For the data vector considered here, this difference is of the same order as the impact of the CDC-effect we described in Sect. 4.2. However, the strength of the latter will most likely increase for non-Gaussian covariances, as we explained at the end of the last section.

## 6. Conclusions

An accurate likelihood analysis plays an essential role in future precision cosmology. We can only exploit the full potential of upcoming high quality data, if we use appropriate statistical methods. In this context the derivation of covariances is an important issue in order not to bias the parameter constraints. In cosmic shear, there are several methods to derive covariances. First, one can calculate **C** analytically assuming a

Gaussian shear field. This assumption breaks down on small angular scales ($< 10$ arcmin), where non-linearities of the matter density field start to become important. Second, covariances can be estimated from ray-tracing simulations. Although computationally more expensive, this method automatically accounts for the non-Gaussianity of the shear field. In both methods the covariance is calculated assuming a specific cosmology. In the first case, this cosmology enters in the power spectrum from which $\mathbf{C}$ is calculated, in the second case we estimate $\mathbf{C}$ from numerical simulations, which are also based on a given cosmology. Past cosmic shear data analyses approximate the covariance to be constant in parameter space, therefore assume that its underlying cosmology does not influence the result of a likelihood analysis significantly.

In this paper we have shown that the covariance matrix depends non-negligibly on its underlying cosmology and that this CDC-effect significantly influences the likelihood contours of parameter constraints. To prove this, we calculate 2500 Gaussian covariance matrices for various parameters of $\Omega_{\mathrm{m}} \in [0.2; 0.4]$ and $\sigma_8 \in [0.6; 1.0]$; all other cosmological parameter are held fixed. Even a change of $\Omega_{\mathrm{m}}$ and $\sigma_8$ within the WMAP5 68% confidence levels has a non-negligible impact on the likelihood contours. Here, the value of $q$ deviates by a factor of 1.84 and this deviation increases to 2.76 if one considers the WMAP5 95% confidence levels. Furthermore, we show that the impact of the CDC-effect depends on survey parameters. Although the likelihood contours become smaller, relatively the CDC-effect becomes more important when increasing the survey size or when decreasing the ratio $\sigma_\epsilon^2 / \bar{n}$. Therefore, a proper treatment becomes more important in the future, for large and deep surveys.

To take cosmology dependent covariances into account we present two methods. First, we perform a likelihood analysis with an adaptive covariance matrix. Here, $\mathbf{C}$ is calculated individually for every point in parameter space, assuming the corresponding parameters as the underlying cosmology. For small surveys this method introduces a bias to the best-fit parameter set, which vanishes when going to larger survey sizes. A disadvantage of this approach is its computational costs. Using the analytic expression for Gaussian covariances is already time-consuming, using ray-tracing covariances to include the non-Gaussianity is not feasible with today's computer power. For the Gaussian case we present a fast and convenient scaling relation to derive covariances on a parameter grid. As a side-effect this approach enhances the constraints on cosmology, for the reason that we now incorporate two cosmology dependent quantities into the likelihood analysis instead of only the mean data vector.

In a strict sense the second method does not account properly for the CDC-effect, however it minimizes the error around the maximum likelihood parameter set ($\boldsymbol{\pi}_{\mathrm{ML}}$). The method consists of two steps, first derive $\boldsymbol{\pi}_{\mathrm{ML}}$ through an iterative process, then derive a ray-tracing covariance with $\boldsymbol{\pi}_{\mathrm{ML}}$ as underlying cosmology and incorporate this in the final likelihood analysis. Here, the approximation of a constant covariance is made, however the error in the posterior probability is minimized in the region of interest; in addition, this ansatz incorporates non-Gaussianity which is non-negligible for future surveys. A drawback is the fact the the starting point of the iteration might bias $\boldsymbol{\pi}_{\mathrm{ML}}$. This must be checked carefully before employing this method, otherwise the approximation of a constant covariance fails around $\boldsymbol{\pi}_{\mathrm{ML}}$.

# References

Anderson, T. W. 2003, An Introduction to Multivariate Statistical Analysis (Wiley-Interscience), 623/624

Bacon, D., Refregier, A., & Ellis, R. 2000, MNRAS, 318, 625

Bartelmann, M. & Schneider, P. 2001, Phys. Rep., 340, 291

Efstathiou, G., Bond, J. R., & White, S. D. M. 1992, MNRAS, 258, 1P

Fu, L., Semboloni, E., Hoekstra, H., et al. 2008, A&A, 479, 9

Hartlap, J., Simon, P., & Schneider, P. 2007, A&A, 464, 399

Hetterscheidt, M., Simon, P., Schirmer, M., et al. 2007, A&A, 468, 859

Hilbert, S., Hartlap, J., White, S. D. M., & Schneider, P. 2008, 2008arXiv0809.5035H

Hoekstra, H., Mellier, Y., van Waerbeke, L., et al. 2006, ApJ, 647, 116

Joachimi, B., Schneider, P., & Eifler, T. 2008, A&A, 477, 43

Kaiser, N. 1998, ApJ, 498, 26

Kaiser, N., Wilson, G., & Luppino, G. A. 2000, 2000astro-ph0003338K

Kendall, M. & Stuart, A. 1979, The Advanced Theory of Statistics (Charles Griffin & Company Limited)

Kilbinger, M. & Schneider, P. 2004, A&A, 413, 465

Kilbinger, M. & Schneider, P. 2005, A&A, 442, 69

Komatsu, E., Dunkley, J., Nolta, M. R., et al. 2008, ArXiv e-prints, 803

Massey, R., Rhodes, J., Leauthaud, A., et al. 2007, ApJS, 172, 239

Schneider, P., van Waerbeke, L., Kilbinger, M., & Mellier, Y. 2002a, A&A, 396, 1

Schneider, P., van Waerbeke, L., & Mellier, Y. 2002b, A&A, 389, 729

Schrabback, T., Erben, T., Simon, P., et al. 2007, A&A, 468, 823

Semboloni, E., Mellier, Y., van Waerbeke, L., et al. 2006, A&A, 452, 51

Semboloni, E., van Waerbeke, L., Heymans, C., et al. 2007, MNRAS, 375, L6

Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629

Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, ApJ, 480, 22

van Waerbeke, L., Mellier, Y., Erben, T., et al. 2000, A&A, 358, 30

van Waerbeke, L., Mellier, Y., & Hoekstra, H. 2005, A&A, 429, 75

Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I., & Bernstein, G. 2000, Nature, 405, 143